Phaseolus vulgaris is nodulated by the symbiovar viciae of several genospecies of *Rhizobium laguerreae* complex in a Spanish region where *Lens culinaris* is the traditionally cultivated legume

José David Flores-Félix[1], Fernando Sánchez-Juanes[2], Paula García-Fraile[1], Angel Valverde[3], Pedro F. Mateos[1], José Manuel Gónzalez-Buitrago[2,4], Encarna Velázquez[1]*, Raúl Rivas[1]

[1]Departamento de Microbiología y Genética and Instituto Hispanoluso de Investigaciones Agrarias (CIALE). Universidad de Salamanca. Edificio Departamental de Biología. Lab 209. Av. Doctores de la Reina S/N. 37007 Salamanca

[2]Instituto de Investigación Biomédica de Salamanca (IBSAL), Complejo Asistencial Universitario de Salamanca, Universidad de Salamanca, CSIC, Salamanca, España

[3]Department of Microbial, Biochemical and Food Biotechnology, University of the Free State, Bloemfontein, South Africa

[4]Departamento de Bioquímica y Biología Molecular, Universidad de Salamanca, Salamanca, España

*Corresponding author: Encarna Velázquez. Departamento de Microbiología y Genética. Lab. 209. Edificio Departamental de Biología. Campus Miguel de Unamuno. 37007 Salamanca. Spain. Phone: +34 923 294532. Fax number: +34 923 294611. E-mail: evp@usal.es

Short Title: *Rhizobium laguerreae sv viciae* nodulates *Phaseolus vulgaris*

Abstract

Phaseolus vulgaris and *Lens culinaris* are two legumes with different distribution centers that were introduced in Spain at different times, but in some regions *L. culinaris* has been traditionally cultivated and *P. vulgaris* did not. Here we analysed the rhizobia isolated from nodules of these two legumes in one of these regions. MALDI-TOF MS analysis showed that all isolated strains matched with *Rhizobium laguerreae* and the phylogenetic analysis of *rrs*, *atpD* and *recA* genes confirmed these results. The phylogenetic analysis of these core genes allowed the differentiation of several groups within *R. laguerreae* and unexpectly, strains with housekeeping genes identical to that of the type strain of *R. laguerreae* presented some differences in the *rrs* gene. In some strains this gene contains an intervening sequence (IVS) identical to that found in *Rhizobium* strains nodulating several legumes in different geographical locations. The *atpD*, *recA* and *nodC* genes of all strains from this study clustered with those of strains nodulating *L. culinaris* in its distribution centers, but not with those nodulating *P. vulgaris* in theirs. Therefore, all these strains belong to the symbiovar viciae, including those isolated from *P. vulgaris* which established symbiosis with the common endosymbiont of *L. culinaris*, instead to with its common endosymbiont, the symbiovar phaseoli. These results are particularly interesting for biogeography studies, because they showed that, due its high promiscuity degree, *P. vulgaris* is able to establish symbiosis with local symbiovars well established in the soil after centuries of cultivation with other legumes.

Keywords: *Rhizobium laguerreae*; symbiovar viciae; *Phaseolus vulgaris*; *Lens culinaris*; Spain

1. Introduction

Phaseolus vulgaris and *Lens culinaris* are two legumes indigenous to Central and South America [13] and to the Middle East [23], respectively. Both legumes coevolved in their respective distribution centers with fast-growing rhizobial strains belonging to genus *Rhizobium*, which are able to nodulate and to fix atmospheric nitrogen in symbiosis with these legumes [39, 40]. The symbiovar phaseoli is the main endosymbiont of *P. vulgaris* in its American distribution centers [3, 4, 7, 14, 37], whereas *L. culinaris* is nodulated by the symbiovar viciae in the Middle East distribution centers of this legume [27, 28].

Both legumes, *P. vulgaris* and *L. culinaris*, currently worldwide cultivated, were introduced in other continents from their distribution centers at different times in history. In Spain, when *P. vulgaris* was introduced, *L. culinaris* was already cultivated for centuries in several regions, and in those with a dominance of rainfed fields it continues to be the traditionally cultivated legume whereas *P. vulgaris* was never introduced. In this work we selected one of these regions where rhizobia have not been analized to date in order to compare the species and symbiovars nodulating *P. vulgaris* and *L. culinaris*. This study is particularly interesting because two strains nodulating *P. vulgaris* in this region have been surprisingly classified into a phylogenetic group comprising *R. laguerreae* [10, 15], a species containing strains of symbiovar viciae nodulating legumes from the cross inoculation group of *Vicia* [5, 6, 30, 35], which has identical *rss* gene, but divergent *recA* and *atpD* genes, that *R. leguminosarum* [30].

Therefore, the first aim of this work was the identification of strains isolated from *P. vulgaris* and *L. culinaris* nodules in a Spanish region where *L. culinaris* has been traditionally cultivated using MALDI-TOF MS (Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry). This methodology was shown to be an useful tool to differentiate among fast-growing rhizobial species [9], and its database has been actualized in this work with the addition of several recently described *Rhizobium* species nodulating *L. culinaris* and *P. vulgaris* in their distribution centers [39, 40]. The second aim of this work was to investigate the phylogenetic relationships among the isolated strains and those nodulating the same hosts in their respective distribution centers through the analysis of the three core genes *rrs, recA* and *atpD* and the symbiotic gene *nodC*, which are commonly used in the phylogenetic analysis of rhizobial species and symbiovars, respectively [24].

Material and Methods

Strains isolation and nodulation tests

All strains were isolated in the region named "La Armuña" (Salamanca, Spain) where *L. culinaris* is commonly cultivated. The rhizobia nodulating *L. culinaris* var. "rubia" were isolated from 10 plants recovered in two fields cultivated with this legume. The rhizobia nodulating *P. vulgaris* var. "pinta" were isolated using this legume as trap plant in soil samples collected in the same region because this legume is not cultivated here. All strains were isolated on YMA plates according to Vincent [42] and reinfection experiments were carried out on the same plants from which they were isolated as was previously described [26].

MALDI-TOF MS performing and data analysis

The sample preparation and the MALDI-TOF MS analysis were carried out as was previously published [9] using a matrix of saturated solution of α-HCCA (Bruker Daltonics, Germany) in 50% acetonitrile and 2.5% trifluoracetic acid. We used amounts of biomass between 5 to 100 mg to obtain the spectra as indicate the manufacturer. The calibration mass were the Bruker Bacterial Test Standards (BTS) which were as follows (masses as averages): RL36, 4365.3 Da; RS22, 5096.8 Da; RL34, 5381.4 Da; RL33meth, 6255.4 Da; RL29, 7274.5 Da; RS19, 10300.1 Da; RNase A, 13683.2 Da and myoglobin, 16952.3 Da.

The score values proposed by the manufacturer are the following: a score value between 2.3 and 3.00 indicates highly probable species identification; a score value between 2.0 and 2.299 indicates secure genus identification and probable species identification, a score value between 1.7 and 1.999 indicates probable genus identification, and a score value <1.7 indicates no reliable identification.

The type strains of recently described species nodulating *P. vulgaris* or *L. culinaris* such as *R. esperanzae*, *R. acidisoli*, *R. hidalgonense*, *R. ecuadorense*, *R. lentis*, *R. binae* and *R. bangladeshense* and *R. laguerreae* FB206^T were added to our database [9]. To add the strains to the reference library, 36 independent spectra were recorded for each strain (three independent measurements at twelve different spots each). Manual/visual estimation of the mass spectra was

performed using Flex Analysis 3.0 (Bruker Daltonics GmbH, Germany) performing smoothing and baseline substraction. Checking existence of flatlines, outliers or single spectra with remarkable peaks differing from the other spectra was done, taking into account that mass deviation within the spectra set shall not be more than 500 ppm. Finally, 20 spectra were selected, removing questionable spectra from the collection. To create peak lists of the spectra, the BioTyper software (Bruker Daltonics GmbH, Germany) was used as described above. The 20 independent peak lists of a strain were used for automated "main spectrum" generation with default settings of the Biotyper software. Thereby, for each library entry a reference peak list (main spectrum) which contains information about averaged masses, averaged intensities, and relative abundances in the 20 measurements for all characteristic peaks of a given strain was created, so a main spectrum displayed the most reproducible peaks typical for a certain bacterial strain.

Cluster analysis was performed based on comparison of strain-specific main spectra created as described above. The dendrogram was constructed by the statistical toolbox of Matlab 7.1 (MathWorks Inc., USA) integrated in the MALDI Biotyper 3.0 software. The parameter settings were: 'Distance Measure=Correlation' and 'Linkage=Average'. The linkage function is normalized according to the distance between 0 (perfect match) and 1000 (no match).

Phylogenetic analyses of rrs, atpD, recA and nodC genes

The amplification and sequencing of *rrs*, *recA* and *atpD*, and *nodC* genes were carried out as indicated by Rivas *et al.* [29], Gaunt *et al.* [12] and Laguerre *et al.* [18], respectively. The sequences obtained were compared with those from the GenBank using the BLASTN program [1]. The obtained sequences and those of related bacteria retrieved from GenBank were aligned using the Clustal W program [36]. The phylogenetic distances were calculated according to Kimura´s two-parameter model [16]. The phylogenetic trees were inferred using the neighbour joining model [31] and MEGA 7.09 [17] was used for all the phylogenetic analyses.

RNA extraction and C-DNA synthesis

Total RNA was extracted from bacterial cells using the MasterPureTM RNA Purification Kit (Epicentre) following manufacturer's instructions. Starting from 1 μg of total RNA,

first-strand cDNA was synthesized by reverse transcription with random hexamers using the ImProm-II™ Reverse Transcription System (Promega). Second-strand synthesis was performed by strand displacement with *Escherichia coli* ligase, DNA polymerase I and RNAse H (all from New England Biolabs). Reverse-transcription negative PCR control reactions (without reverse transcriptase) were run to check for DNA contamination.

Results and discussion

MALDI-TOF MS analysis

MALDI-TOF MS is a reliable method for bacterial identification originally conceived for the identification of pathogenic bacteria [2, 38]. Nevertheless, this methodology has proven to be useful for the identification of other bacterial such as those from family *Rhizobiaceae* [9]. The results obtained by MALDI-TOF MS analysis showed that all strains isolated from both *P. vulgaris* and *L. culinaris* nodules matched with the type strain of *R. laguerreae* FB206^T with score values higher than 2.0, including the two strains previously isolated from *P. vulgaris* in the same region, PEPV16 and PEPV40 [10, 15] (Table 1). The mathematical analysis of the spectra from the strains of this study, showed that they are divided into seven groups with internal distance levels equal or lower than 0.8 (Fig. S1, Table I). Three of these groups contain strains nodulating *P. vulgaris* and *L. culinaris*, whereas the remaining groups only contain strains nodulating one of the two legumes (Fig. S1). Representative strains from these groups were selected for the phylogenetic analysis of core and symbiotic genes.

Analysis of the rrs gene

The *rrs* genes of the type strains from several species within the phylogenetic group of *R. leguminosarum*, including the species *R. laguerreae*, are identical (Fig. 1). Nevertheless, slight variations were found in the *rrs* genes of some strains belonging to this phylogenetic group isolated in Spain [41]. Therefore, we analysed this gene in the strains isolated in this work (Fig. 1). The results of the analysis in representative strains from MALDI-TOF MS groups showed that they belong to five groups with some differences in their *rrs* genes (Fig. 1). The *rrs* gene sequence of strain PEPV11 was identical to that of *R. laguerreae*

FB206^T and also to the strains from the *rrs* type 1 of Villadas *et al.* [41] represented in the phylogenetic tree by the strain li10. All these sequences also were 100% identical to those of the type strains of *R. leguminosarum*, *R. indigoferae* , *R. anhuiense*, *R. acidisoli*, *R. hidalgonense* and *R. sophorae* (cluster I). The *rrs* gene of strain PEPV16 was identical to those of strains presenting the *rrs* type 2 of Villadas *et al.* [41] represented in the tree by the strain lc10 (cluster II). These strains have a different nucleotide in their *rrs* genes with respect to *R. laguerreae* FB206^T (in the position 1070 taking the accession number JN558651 as reference). The *rrs* type 3 from Villadas *et al.* [41], represented in the phylogenetic tree by the strain vd48, was not found in any of the strains from this study (lineage VI). The strains MLS17, PEPV31 and PEPV37 (cluster III) showed 2 different nucleotides with respect to *R. laguerreae* FB206^T (in the positions 942 and 955 taking the accession number JN558651 as reference). The strains isolated from *L. culinaris* in the distribution centers of Syria and Turkey by Rashid *et al.* [27] have two different nucleotides in the positions 942 and 955 taking the *rrs* accession number JN558651 as reference (data not shown), but they were not included in Fig. 1 because their *rrs* gene sequences are not complete. The strains MLSC02 and PEPV40 (cluster IV) presented 3 different nucleotides (in the positions 940, 942 and 955 taking the accession number JN558651 as reference). Finally, the *rrs* genes of strains PEPV08, MLSC04 and MLS05 (cluster V) have 4 different nucleotides with respect to *R. laguerreae* FB206^T (in the positions 67, 69, 942 and 955 taking the accession number JN558651 as reference) and they also contain an intervening sequence (IVS) of 73 bp located at the begining of this gene (between the nucleotides 70 and 143 taking the accession number JN558651 as reference). A similar IVS was reported by Willems and Collins (1993) in the *rrs* gene of *Rhizobium leucaenae* CFN 299^T, previously named *Rhizobium tropici* IIA. The absence of this IVS in the transcribed rRNA of these four strains was demonstrated by comparing the amplicons corresponding to the *rrs* gene derived from DNA and RNA. All four isolates showed identical *rrs* gene sequences and sizes (data not shown) as was previously reported for other bacteria [25].

A search in Genbank showed that IVS are widely distributed in *rrs* genes of strains from the order *Rhizobiales*, classified into genera *Rhizobium*, *Sinorhizobium* (currently *Ensifer*) and *Agrobacterium* belonging to family *Rhizobiaceae*, *Mesorhizobium* from family *Phyllobacteriaceae* and *Ochrobactrum* from family *Brucellaceae* (Fig. 2A). The IVS of these strains are placed at the same position within the *rrs* gene and have identical

sequences at the begining (7 nt) and at the end (8 nt). Nevertheless, despite the low size of this fragment, their sequences are different in the strains analysed to date.

Analysing in parallel the IVS and the *rrs* genes of strains whose complete sequences (more than 1400 nt) are available in Genbank (Fig. 2A and 2B) we found that several of these strains that have been assigned to the genus *Rhizobium* should be included into the new genera *Neorhizobium*, *Pararhizobium* or *Allorhizobium* [21, 22] (Fig. 2B). Some of these strains are also misclassified at species level, since they are named *Rhizobium galegae* and *Rhizobium huautlense* (currently *Neorhizobium galegae* and *Neorhizobium huautlense*), although their *rrs* genes are divergent to those from the type strains of *N. galegae* and *N. huautlense.* Other strains named *R. leguminosarum* are also misclassified at species level, since their *rrs* genes are phylogenetically divergent of the type strain of this species and the strains named *Rhizobium tropici* should be renamed as *Rhizobium leucaenae* (Fig. 2B). Concerning to the strains isolated from *L. culinaris* in its distribution centers (Syria and Turkey) by Rashid *et al.* [27], we can not include them in the phylogenetic analyses since it lacks the region where the IVS is located in their *rrs* gene sequences.

The analysis of the IVS showed high variability among the members of order *Rhizobiales* particularly within genus *Rhizobium*, which to date contains most of strains harbouring IVS in their *rrs* genes (Fig. 2A). It has been proposed that the IVS from *rrs* genes can be useful as a phylogenetic marker at species or strain levels [25]. For example, the IVS of *Faecalibacterium* has a broad geographic distribution and it can be used to detect fecal pollution in water [8, 32, 34] and it has been reported that *rrs* IVSs can be specific for faecal bacteria from different hosts and can be used as genetic markers for microbial source tracking [33]. The strains from order *Rhizobiales* have several types of IVS with a variable divergence degree among them and we did not found a clear relationship between the IVS and a taxonomic affiliation (Fig. 2A). In fact, strains belonging to the same genus (i.e.: *Rhizobium*, *Pararhizobim* and *Neorhizobium*) showed divergent IVS types, whereas strains from different genera (i.e.: *Mesorhizobium* and *Ochrobactrum*) showed the same IVS type (Fig. 2A). Nevertheless, the analysis of IVS could be useful for biogeography studies, since in some strains identical IVS are linked to identical *rrs* genes indicating a geographic dispersion of the same species/strains. The best example are the strains of *R. leucaenae* which have identical *rrs* genes, including IVS, with independence of their isolation site, America or Europe, suggesting a common origin for all of them probably due to the dispersion of these strains together with the seeds. However, from our results we

cannot conclude about the usefulness of the IVS as a phylogenetic marker of genera or species for the order *Rhizobiales*, as strains from different species and even genera belonging to this order have identical IVS.

The most relevant finding from the *rrs* gene analysis in this work was the existence of strains with different sequences in this gene, two of them identical to those of strains previously isolated in South Spain [41], but others found by first time in this study. Since all of them belong to the phylogenetic group of *R. leguminosarum*, which contains several species with identical *rrs* gene sequences and phylogenetically divergent *recA* and *atpD* genes, the taxonomic affiliation of the strains from this work should be clarified on the basis of the analysis of these genes.

Analysis of recA and atpD genes

The results of the phylogenetic analysis of *recA* and *atpD* genes showed that, with the exception of strains PEPV40 and MLSC02, which have *recA* and *atpD* genes identical to those of *R. laguerreae* FB206^T, the remaining strains showed similarity values lower than 96% and 99%, respectively. Nevertheless, considering both genes together, the closest related species to our strains is *R. laguerreae* in agreement with the results of MALDI-TOF MS.

The analysis of the *recA* and *atpD* genes of the strains isolated in this study showed the existence of divergence in these genes among strains having identical *rrs* genes and *vice versa*. For example, within the cluster A, the strain PEPV11 presented a *rrs* gene identical to that of *R. laguerreae* FB206^T (*rrs* cluster I) but different *recA* and *atpD* genes. The strains PEPV31 (*rrs* cluster III), PEPV40 and MLSC02 (*rrs* cluster IV) have different *rrs* genes, but identical *recA* and *atpD* genes that *R. laguerreae* FB206^T (Fig. 3). The strains MLS17 and PEPV37 with identical *rrs* gene that the strain PEPV31 (*rrs* cluster III) also belong to different clusters in the housekeeping genes analysis, B and C, respectively. To this last cluster C also belong the strains PEPV16 (*rrs* cluster II) and the strains MLS05, MLSC04 and PEPV08 which have an IVS in their *rrs* genes and belong to *rrs* cluster IV. Therefore, the strains from clusters B and C belong to two genospecies of *R. laguerreae*, which can be differentiated from the type strain of this species on the basis of both *rrs* and housekeeping genes.

In the analysis of the concatenated *recA* and *atpD* genes we included representative strains of strains isolated in *L. culinaris* and *P. vulgaris* distribution centers (Fig. 3). The results showed that all strains isolated in this study were phylogenetically related to strains isolated from *L. culinaris* nodules in its distribution centers, Syria and Turkey, and that none of them were close to the species nodulating *P. vulgaris* in its American distribution centers (Fig. 3). The strains isolated from *L. culinaris* in Syria and Turkey are included in the three clusters, A, B and C, occupied by the strains isolated in this study and additionally two of these strains formed the fourth cluster D. Since all these strains were more closely related to *R. laguerreae* than the remaining species of genus *Rhizobium* their current name *R. leguminosarum* should be changed by *R. laguerreae* (Fig. 3). Three of these strains isolated in Turkey, TLR9, TLR10 and TLR14, presented identical sequences of *recA* and *atpD* genes than some of Spanish strains indicating a distribution of these strains from this distribution center to other geographical locations (Fig. 3). The remaining strains, although they are phylogenetically related with Spanish strains, formed several divergent lineages among them and with respect to these strains (Fig. 3). The clusters containing *L. culinaris* and *P. vulgaris* nodulating strains isolated in this work do not contained strains isolated from *P. vulgaris* in its distribution centers from America, including two strains initially classified in the species *R. leguminosarum*, NH05 and CCGM1 which currently belong to *R. hidalgonense* and *R. phaseoli*, respectively (Fig. 3). The remaining strains isolated from *P. vulgaris* nodules in American countries are distributed in different clusters or lineages closely related to already described American species nodulating this legume.

Therefore, in agreement with the results of the phylogenetic analysis of *rrs* gene, those from that of *recA* and *atpD* genes support that several genospecies within *R. laguerreae* are the common endosymbionts of *L. culinaris* in several geographical locations, including its distribution centers, and also of promicuous legumes such as *P. vulgaris* when they are cultivated in soils where *L. culinaris* is the common cultivated legume.

nodC gene analysis

The *nodC* gene is currently the main phylogenetic marker for *Rhizobium* symbiovars [24] and the phylogenetic analysis of this gene allowed the differentiation of symbiovars phaseoli and viciae nodulating *P. vulgaris* and *L. culinaris*, respectively, in their

distribution centers (Fig. 4). The results of the *nodC* gene analysis showed that all strains isolated in this study belong to the symbiovar viciae with independence of the legume from which they were isolated (Fig. 4). Nevertheless, these strains belong to three different clusters (1, 2 and 3), all of them encompasing strains isolated from both *L. culinaris* and *P. vulgaris* in this work. The cluster 1 encompased the type strains of *R. laguerreae*, *R. fabae* and *R. binae*, two strains isolated from *L. culinaris* and *P. vulgaris* in this work and several strains isolated from *L. culinaris* in Syria and Turkey (Fig. 4). The cluster 2 contains three strains isolated from *L. culinaris* and *P. vulgaris* in this work and three strains isolated in Turkey. The cluster 3 contains five strains isolated from *L. culinaris* and *P. vulgaris* in this work and reference strains for *R. lentis* and *R. bangladeshense.* Finally, a strain isolated in Syria from *L. culinaris* nodules formed an independent lineage phylogenetically related with the cluster 3 (Fig. 4). The strains which have housekeeping genes identical to those of strains TLR9, TLR10 and TLR14 do not carried identical *nodC* genes showing a high diversification degree of this gene in different geographical locations. Any of the strains nodulating *P. vulgaris* in this study carry *nodC* genes phylogenetically related to those carried by strains of symbiovar phaseoli isolated in American and European countries. This confirms that the symbiovar phaseoli has an American origin and that their symbiotic genes were transferred to other species nodulating *P. vulgaris* in regions where it has been commonly cultivated [4, 11, 37]. In soils where this legume has been not cultivated, as occurs in this case, the high promiscuity degree of this legume [19, 20] allows the establishment of symbiosis with rhizobia nodulating local legumes.

In summary, the results of this work showed that the American legume *P. vulgaris* establish symbiosis with the symbiovar viciae of several genospecies of *R. laguerreae* complex in a region from Northwest Spain where *L. culinaris* is the traditionally cultivated legume. The phylogenetic analysis of three core genes allowed the differentiation of several groups within *R. laguerreae* and unexpectly, strains with housekeeping genes identical to that of the type strain of *R. laguerreae* presented some differences in the *rrs* gene. The *rrs* genes of some strains have an intervening sequence (IVS) identical to that found in *Rhizobium* strains nodulating several legumes in different geographical locations. The *atpD*, *recA* and *nodC* genes of all strains clustered with those of strains nodulating *L. culinaris* in its distribution centers, but not with those nodulating *P. vulgaris* in theirs. The nodulation of *P. vulgaris* in regions where it has not been previously cultivated is more likely due to the high promiscuity degree of this legume, which can nodulate with several

symbiovars, the symbiovar viciae in this case. These findings showed the need of performing further studies of strains nodulating *P. vulgaris* in different European soils where other legumes than *P. vulgaris* are commonly cultivated in order to have a more complete picture of *Rhizobium*-legume symbiosis.

Aknowledgements

This work was supported by JCyL (Junta de Castilla y León, Spanish Regional Government) and MICINN (Ministry of Science and Innovation, Spanish Central Government). José David Flores-Félix is recipient of a predoctoral fellowship from University of Salamanca. José David Flores-Félix is recipient of a postdoctoral Marie Skłodowska-Curie fellowship.

References

[1] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

[2] Angeletti, S. (2017) Matrix assisted laser desorption time of flight mass spectrometry (MALDI-TOF MS) in clinical microbiology. J. Microbiol. Meth. 138, 20–29.

[3] Aguilar, O.M., Riva, O., Peltzer, E. (2004) Analysis of *Rhizobium etli* and of its symbiosis with wild *Phaseolus vulgaris* supports coevolution in centers of host diversification. Proc. Nat. Acad. Sci. USA 101, 13548–13553.

[4] Amarger, N., Martínez-Romero, E., Olivares, J., Sanjuan, J. (1999) At least five rhizobial species nodulate *Phaseolus vulgaris* in a Spanish soil. FEMS Microbiol. Ecol. 30, 87–97.

[5] Belhadi, D., de Lajudie, P., Ramdani, N., Le Roux, C., Boulila, F., Tisseyre, P., Boulila, A., Benguedouar, A., Kaci, Y., Laguerre, G. (2018) *Vicia faba* L. in the Bejaia region of Algeria is nodulated by *Rhizobium leguminosarum* sv. viciae, *Rhizobium laguerreae* and two new genospecies. Syst, Appl. Microbiol. 41, 122–130.

[6] Benidire, L., Lahrouni, M., Daoui, K., Fatemi, Z.E.A., Gómez Carmona, R., Göttfert, M., Oufdou, K. (2017) Phenotypic and genetic diversity of Moroccan rhizobia isolated

366 from *Vicia faba* and study of genes that are likely to be involved in their osmotolerance.
367 Syst. Appl. Microbiol. 41, 51–61.

368 [7] Díaz-Alcántara, C.A., Ramírez-Bahena, M.H., Mulas, D., García-Fraile, P., Gómez-
369 Moriano, A., Peix, A., Velázquez, E., González-Andrés, F. (2014) Analysis of rhizobial
370 strains nodulating *Phaseolus vulgaris* from Hispaniola Island, a geographic bridge between
371 Meso and South America and the first historical link with Europe. Syst. Appl. Microbiol.
372 37, 149–156.

373 [8] Duan, C., Cui, Y., Zhao, Y., Zhai, J., Zhang, B., Zhang, K., Sun, D., Chen H. (2016)
374 Evaluation of *Faecalibacterium* 16S rDNA genetic markers for accurate identification of
375 swine faecal waste by quantitative PCR. J. Environ. Management 181, 193–200.

376 [9] Ferreira, L., Sánchez-Juanes, F., García-Fraile, P., Rivas, R., Mateos, P.F., Martínez-
377 Molina, E., González-Buitrago, J.M., Velázquez, E. (2011) MALDI-TOF mass
378 spectrometry is a fast and reliable platform for identification and ecological studies of
379 species from family *Rhizobiaceae*. PLoS One 6, e20223.

380 [10] Flores-Félix, J.D., Marcos-García, M., Silva, L.R., Menéndez, E., Martínez-Molina,
381 E., Mateos, P.F., Velázquez, E., García-Fraile, P., Andrade, P., Rivas, R. (2015)
382 Rhizobium as plant probiotic for strawberry production under microcosm conditions.
383 Symbiosis 67, 25–32.

384 [11] García-Fraile, P., Mulas-García, D., Rivas, R., González-Andrés, F., Velázquez, E.
385 (2010) *Phaseolus vulgaris* is nodulated in Northern Spain by *Rhizobium leguminosarum*
386 strains harboring two nodC alleles present in American *Rhizobium etli* strains:
387 biogeographical and evolutionary implications. Can. J. Microbiol. 56, 657–666.

388 [12] Gaunt, M.W., Turner, S.L., Rigottier-Gois, L., Lloyd-Macgilp, S.A., Young, J.W.P.
389 (2001) Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification
390 of rhizobia. Int. J. Syst. Evol. Microbiol. 51, 2037–2048.

391 [13] Gepts, P., Debouck, D. (1991) Origin, domestication, and evolution of the common
392 bean (*Phaseolus vulgaris* L.). In: van Schoonhoven, A., Voysest, O. Common Beans:
393 Research for Crop Improvement, CAB International,UK, pp. 7–55.

394 [14] Grange, L., Hungria, M., Graham, P.H., Martínez-Romero, E. (2007) New insights
395 into the origins and evolution of rhizobia that nodulate common bean (*Phaseolus vulgaris*)
396 in Brazil. Soil Biol. Biochem. 39: 867–876.

[15] Jiménez-Gómez, A., Flores-Félix, J.D., García-Fraile, P., Mateos, P.F., Menéndez, E., Velázquez, E., Rivas, R. (2018) Probiotic activities of *Rhizobium laguerreae* on growth and quality of spinach. Sci. Rep. 8, 295.

[16] Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

[17] Kumar, S., Stecher, G., Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 3, 1870–1874.

[18] Laguerre, G., Nour, S.M., Macheret, V., Sanjuan, J., Drouin, P., Amarger, N. (2001) Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. Microbiology 147, 981–993.

[19] Martínez-Romero, E. (2003) Diversity of *Rhizobium-Phaseolus vulgaris* symbiosis: overview and perspectives. Plant Soil 252, 11–23.

[20] Michiels, J., Dombrecht, B., Vermeiren, N., Xi, C., Luyten, E., Vanderleyden, J. (1998) *Phaseolus vulgaris* is a non-selective host for nodulation. FEMS Microbiol. Ecol., 26, 193–205.

[21] Mousavi SA, Österman J, Wahlberg N, Nesme X, Lavire C, Vial L, Paulin L, de Lajudie P, Lindström K. Phylogeny of the *Rhizobium-Allorhizobium-Agrobacterium* clade supports the delineation of *Neorhizobium* gen. nov. Syst. Appl. Microbiol. 37, 208–215.

[22] Mousavi, S.A., Willems, A., Nesme, X., de Lajudie, P., Lindström, K. (2015) Revised phylogeny of *Rhizobiaceae*: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations. Syst. Appl. Microbiol. 38, 84-90.

[23] Muehlbauer, F.J., McPhee, K.E. (2005) Lentil (*Lens culinaris* Medik.). In: Singh, R.J., Jauhar P.P. (Eds.), Genetic Resources, Chromosome Engineering, and Crop Improvement: Grain Legumes, Volume I, CRC Press Boca Raton, Taylor & Francis group, Florida, USA, pp. 219–228.

[24] Peix, A., Ramírez-Bahena, M.H., Velázquez, E., Bedmar, E.J. (2015) Bacterial Associations with Legumes. Crit. Rev. Plant Sci. 34, 17–42.

[25] Rainey, F.A, Ward-Rainey, N.L., Janssen, P.H., Hippe, H., Stackebrandt, E. (1996) *Clostridium paradoxum* DSM 7308^T contains multiple 16S rRNA genes with heterogeneous intervening sequences. Microbiology 142, 2087–2095.

[26] Ramírez-Bahena, M.H., Velázquez, E., Fernández-Santos, F., Peix, A., Martínez-Molina, E., Mateos, P.F. (2009) Phenotypic, genotypic and symbiotic diversity in strains nodulating clover in different soils in Spain. Can. J. Microbiol. 55, 1207–1216.

[27] Rashid, M.H., Gonzalez, J., Young, J.P., Wink, M. (2014) *Rhizobium leguminosarum* is the symbiont of lentils in the Middle East and Europe but not in Bangladesh. FEMS Microbiol Ecol. 87, 64–77.

[28] Rashid, M.H., Young, J.P., Everall, I., Clercx, P., Willems, A., Santhosh Braun, M., Wink, M. (2015) Average nucleotide identity of genome sequences supports the description of *Rhizobium lentis* sp. nov., *Rhizobium bangladeshense* sp. nov. and *Rhizobium binae* sp. nov. from lentil (*Lens culinaris*) nodules. Int. J. Syst. Evol. Microbiol. 65, 3037–3045.

[29] Rivas, R., García-Fraile, P., Mateos, P.F., Martínez-Molina, E., Velázquez, E. (2007). Characterization of xylanolytic bacteria present in the bract phyllosphere of the date palm *Phoenix dactylifera*. Lett. Appl. Microbiol. 44, 181–187.

[30] Saïdi, S., Ramírez-Bahena, M.H., Santillana, N., Zúñiga, D., Álvarez-Martínez, E., Peix, A., Mhamdi, R., Velázquez, E. (2014) *Rhizobium laguerreae* sp. nov. nodulates *Vicia faba* on several continents. Int. J. Syst. Evol. Microbiol. 64, 242–247.

[31] Saitou, N., Nei, M. (1987) A neighbour-joining method: a new method for reconstructing phylogenetics trees. Mol. Biol. Evol. 44, 406–425.

[32] Shen, Z., Duan, C., Zhang, C., Carson, A., Xu, D., Zheng, G. (2013) Using an intervening sequence of *Faecalibacterium* 16S rDNA to identify poultry feces. Water Res. 47, 6415–6422.

[33] Shen, Z., Zhang, N., Mustapha, A., Lin, M., Xu, D., Deng, D., Reed, M., Zheng, G. (2016). Identification of host-specific genetic markers within 16s rdna intervening sequences of 73 genera of fecal bacteria. J. Data Mining Genomics Proteomics 7, 1.

[34] Sun, D., Duan, C., Shang, Y., Ma, Y., Tan, L., Zhai, J., Gao, X., Guo, J., Wang, G. (2016) Application of *Faecalibacterium* 16S rDNA genetic marker for accurate identification of duck faeces. Environ. Sci. Poll. Res. 23, 7639–7647.

[35] Taha, K., Berraho, E.B., El Attar, I., Dekkiche, S., Aurag, J., Béna, G3. (2017) *Rhizobium laguerreae* is the main nitrogen-fixing symbiont of cultivated lentil (*Lens culinaris*) in Morocco. Syst. Appl. Microbiol. 41, 113–121.

[36] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G. (1997) The clustalX windows interface: flexible strategies for multiple sequence alignement aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

[37] Valverde, A., Igual, J.M., Peix, A., Cervantes, E., Velázquez, E. (2006) *Rhizobium lusitanum* sp. nov. a bacterium that nodulates *Phaseolus vulgaris*. Int. J. Syst. Evol. Microbiol. 56, 2631–2637.

[38] van Belkum, A., Welker, M., Pincus, D., Charrier, J.P,, Girard, V. (2017) Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry in Clinical Microbiology: What Are the Current Issues?. Ann. Lab. Med. 37, 475–483.

[39] Velázquez, E., Carro, L., Flores-Félix, J.D., Martínez-Hidalgo, P., Menéndez, E., Ramírez-Bahena, M.H., Mulas, R., González-Andrés, P., Martínez-Molina, E., Peix A. (2017) The legume nodule microbiome: a source of plant growth-promoting bacteria. In: Kumar V., Kumar M., Sharma S., Prasad R. (Eds). Probiotics and Plant Health. Springer, Singapore. pp 41–70.

[40] Velázquez, E., García-Fraile, P., Ramírez-Bahena, M.H., Rivas, R., Martínez-Molina, E. (2017) Current status of the taxonomy of bacteria able to establish nitrogen-fixing legume symbiosis. In: Zaidi A., Khan M., Musarrat J. (Eds), Microbes for Legume Improvement, Springer, Cham, pp 1–43.

[41] Villadas, P.J., Lasa, A.V., Martínez-Hidalgo, P., Flores-Félix, J.D., Martínez-Molina, E., Toro, N., Velázquez, E., Fernández-López, M. (2017) Analysis of rhizobial endosymbionts of *Vicia*, *Lathyrus* and *Trifolium* species used to maintain mountain firewalls in Sierra Nevada National Park (South Spain). Syst. Appl. Microbiol. 40, 92–101.

[42] Vincent, J.M. (1970). A Manual for the Practical Study of the Root-nodule Bacteria IBP Handbook 15. Black Well Scientific Publications, Oxford.

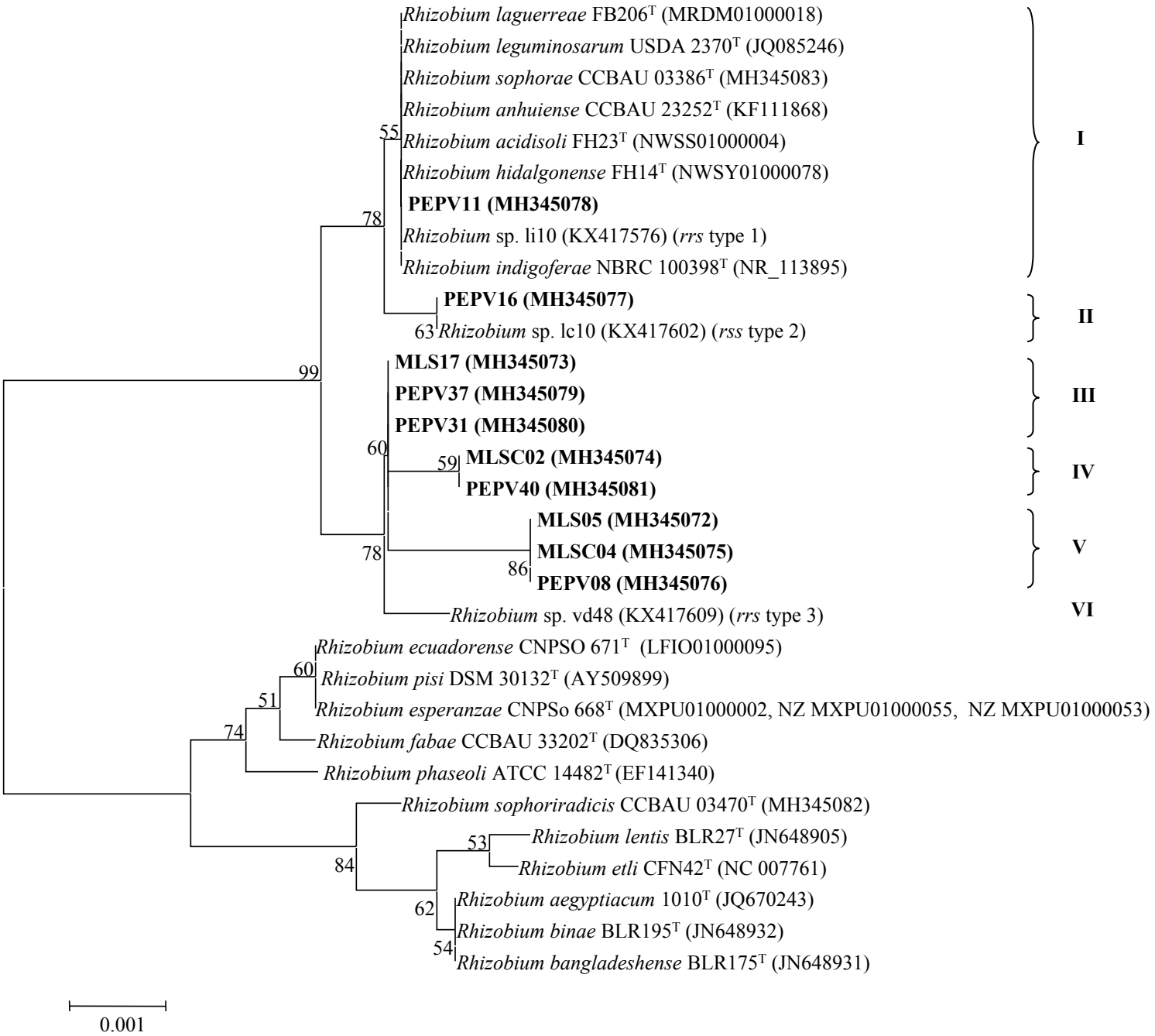
Legends to figures

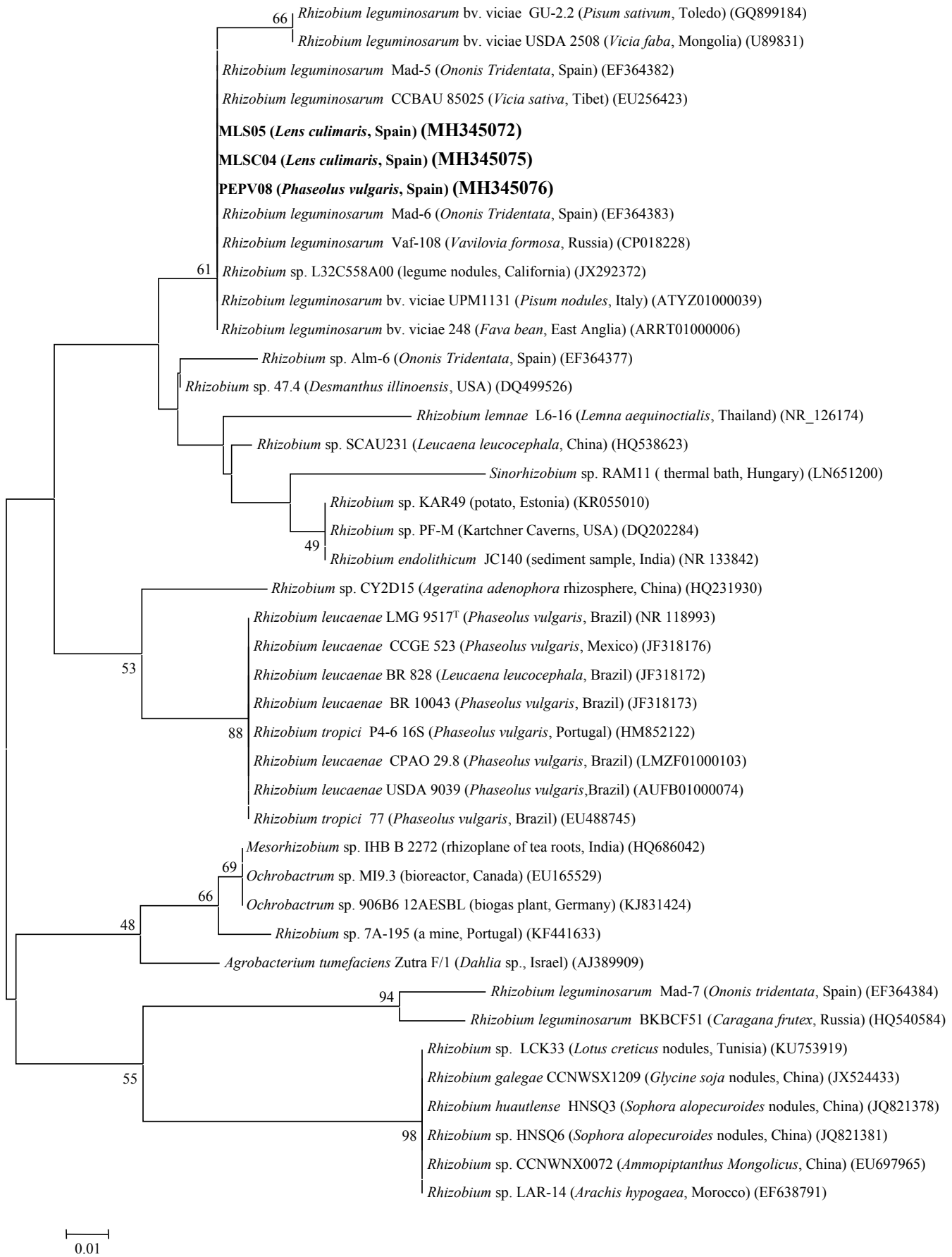
Figure 1. Neighbour-joining phylogenetic rooted tree based on *rrs* gene sequences (1433 nt) showing the taxonomic location of representative strains from different groups of MALDI-TOF MS within the genus *Rhizobium*. Bootstrap values calculated for 1000 replications are indicated. Bar, 1 nt substitution per 1000 nt. Accession numbers from Genbank are given in brackets.

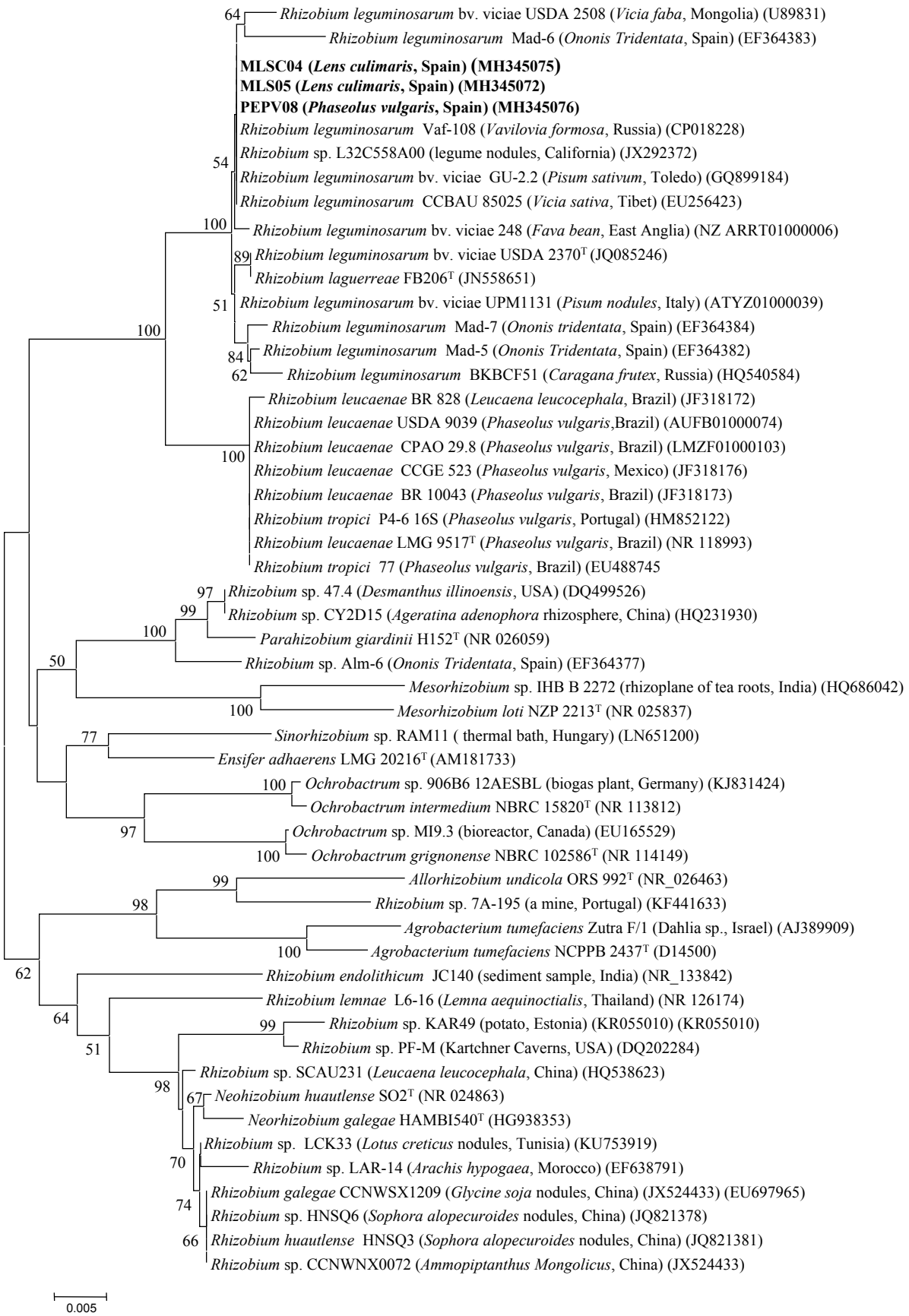
Figure 2. **A)** Neighbour-joining phylogenetic unrooted tree based on IVS sequences (70 nt) of representative strains of this study and those of different species and genera held in Genbank. Bootstrap values calculated for 1000 replications are indicated. Bar, 1 nt substitution per 100 nt. **B)** Neighbour-joining phylogenetic rooted tree based on *rrs* gene sequences (1503 nt) of the same strains. Bootstrap values calculated for 1000 replications are indicated. Bar, 5 nt substitution per 1000 nt. Accession numbers from Genbank are given in brackets.

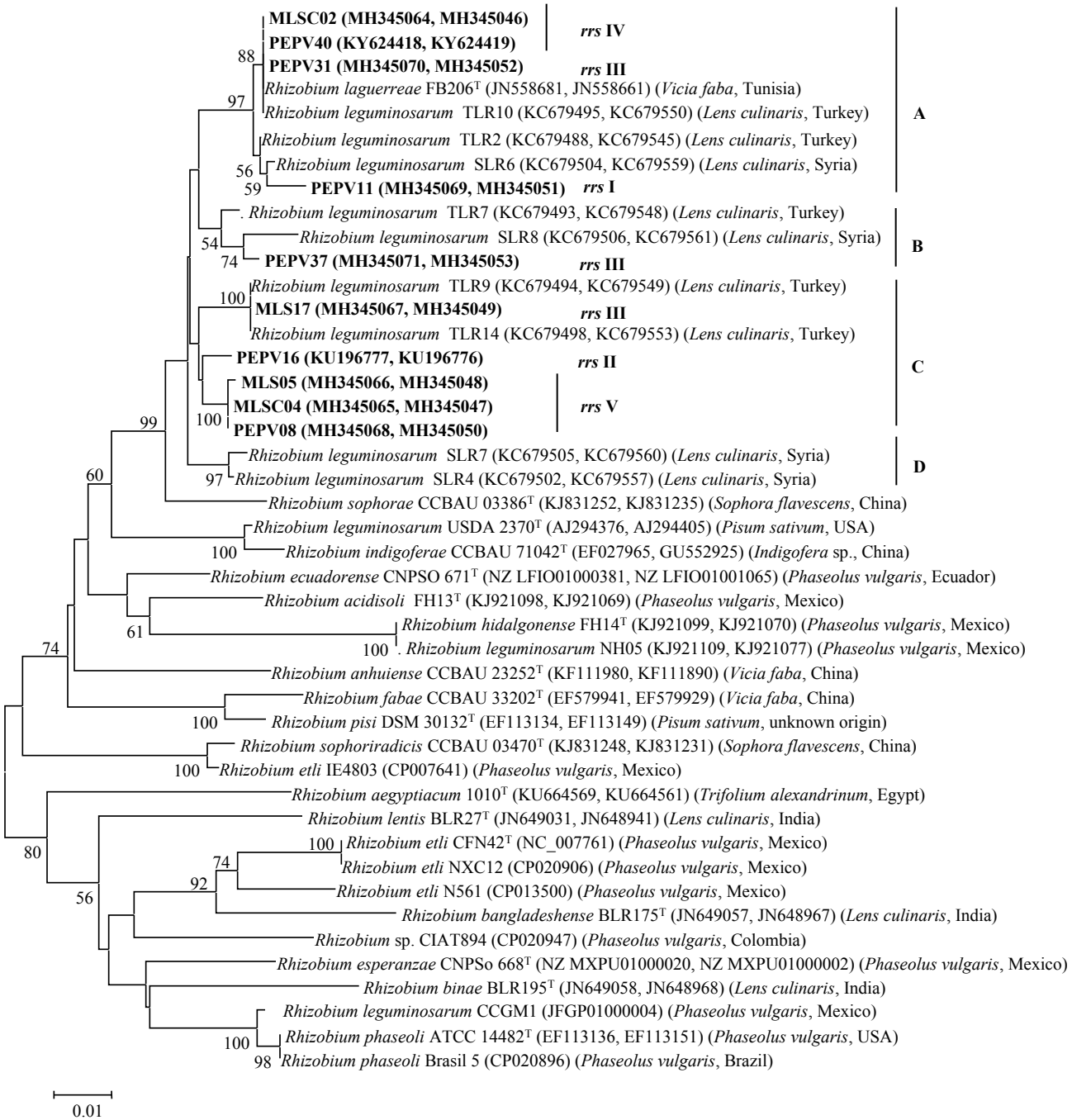
Figure 3. Neighbour-joining phylogenetic tree based on *recA* and *atpD* concatenated gene sequences (728 nt) showing the position of representative strains from each group within genus *Rhizobium*. Bootstrap values calculated for 1000 replications are indicated. Bar, 1 nt substitution per 100 nt. Accession numbers from Genbank are given in brackets.

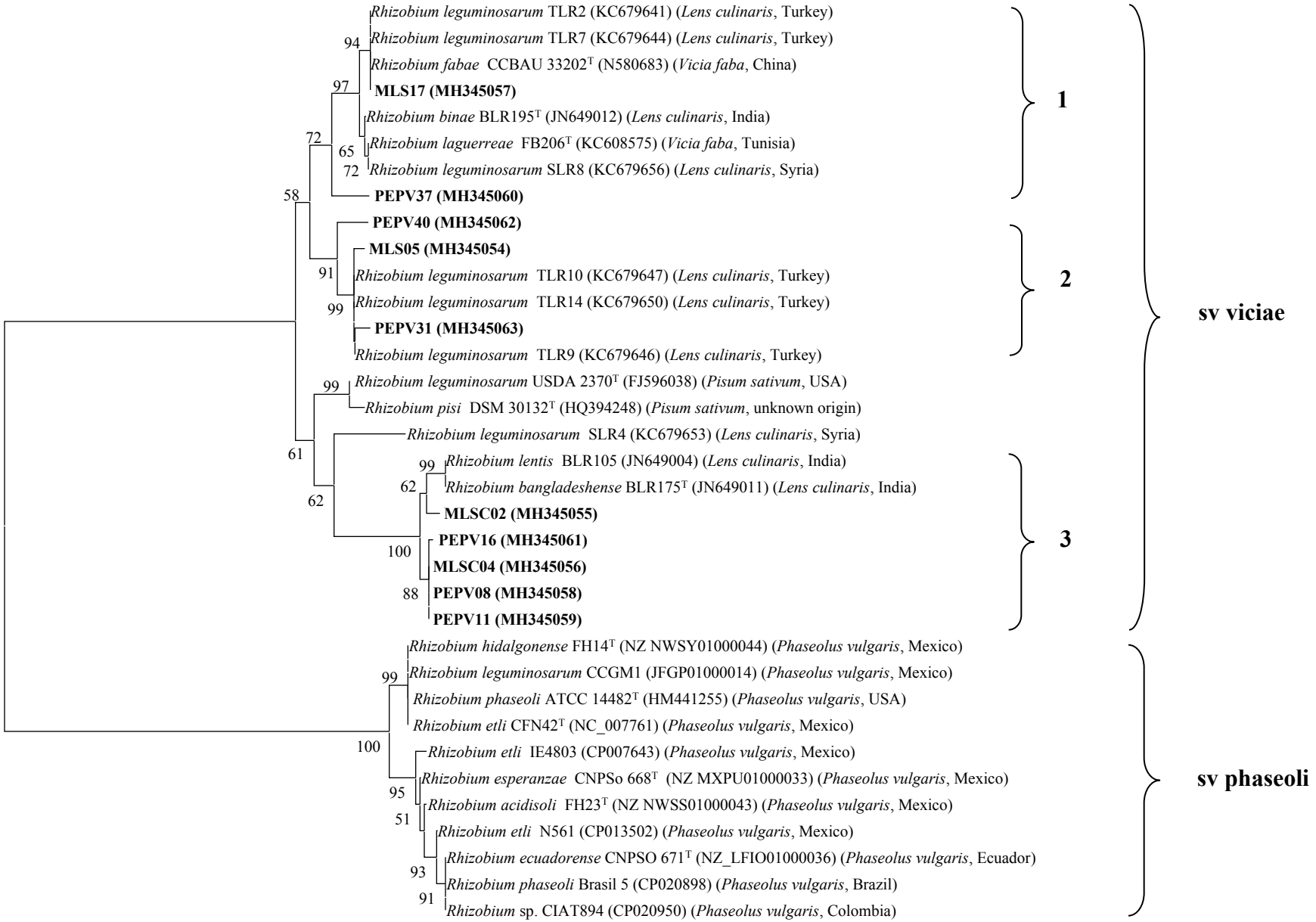
Figure 4. Neighbour-joining phylogenetic tree based on *nodC* gene sequences (730 nt) showing the position of representative strains from each group within the symbiovars viciae and phaseoli. Bootstrap values calculated for 1000 replications are indicated. Bar, 5 nt substitution per 100 nt. Accession numbers from Genbank are given in brackets.


Rhizobium laguerreae FB206^T (MRDM01000018)
Rhizobium leguminosarum USDA 2370^T (JQ085246)
Rhizobium sophorae CCBAU 03386^T (MH345083)
Rhizobium anhuiense CCBAU 23252^T (KF111868)
Rhizobium acidisoli FH23^T (NWSS01000004)
Rhizobium hidalgonense FH14^T (NWSY01000078)
PEPV11 (MH345078)
Rhizobium sp. li10 (KX417576) (*rrs* type 1)
Rhizobium indigoferae NBRC 100398^T (NR_113895)
PEPV16 (MH345077)
Rhizobium sp. lc10 (KX417602) (*rss* type 2)
MLS17 (MH345073)
PEPV37 (MH345079)
PEPV31 (MH345080)
MLSC02 (MH345074)
PEPV40 (MH345081)
MLS05 (MH345072)
MLSC04 (MH345075)
PEPV08 (MH345076)
Rhizobium sp. vd48 (KX417609) (*rrs* type 3)
Rhizobium ecuadorense CNPSO 671^T (LFIO01000095)
Rhizobium pisi DSM 30132^T (AY509899)
Rhizobium esperanzae CNPSo 668^T (MXPU01000002, NZ MXPU01000055, NZ MXPU01000053)
Rhizobium fabae CCBAU 33202^T (DQ835306)
Rhizobium phaseoli ATCC 14482^T (EF141340)
Rhizobium sophoriradicis CCBAU 03470^T (MH345082)
Rhizobium lentis BLR27^T (JN648905)
Rhizobium etli CFN42^T (NC 007761)
Rhizobium aegyptiacum 1010^T (JQ670243)
Rhizobium binae BLR195^T (JN648932)
Rhizobium bangladeshense BLR175^T (JN648931)
55
78
63
99
60
59
78
86
60
51
74
53
84
62
54
I
II
III
IV
V
VI
0.001


66
Rhizobium leguminosarum bv. viciae GU-2.2 (Pisum sativum, Toledo) (GQ899184)
Rhizobium leguminosarum bv. viciae USDA 2508 (Vicia faba, Mongolia) (U89831)
Rhizobium leguminosarum Mad-5 (Ononis Tridentata, Spain) (EF364382)
Rhizobium leguminosarum CCBAU 85025 (Vicia sativa, Tibet) (EU256423)
MLS05 (Lens culimaris, Spain) (MH345072)
MLSC04 (Lens culimaris, Spain) (MH345075)
PEPV08 (Phaseolus vulgaris, Spain) (MH345076)
Rhizobium leguminosarum Mad-6 (Ononis Tridentata, Spain) (EF364383)
Rhizobium leguminosarum Vaf-108 (Vavilovia formosa, Russia) (CP018228)
61
Rhizobium sp. L32C558A00 (legume nodules, California) (JX292372)
Rhizobium leguminosarum bv. viciae UPM1131 (Pisum nodules, Italy) (ATYZ01000039)
Rhizobium leguminosarum bv. viciae 248 (Fava bean, East Anglia) (ARRT01000006)
Rhizobium sp. Alm-6 (Ononis Tridentata, Spain) (EF364377)
Rhizobium sp. 47.4 (Desmanthus illinoensis, USA) (DQ499526)
Rhizobium lemnae L6-16 (Lemna aequinoctialis, Thailand) (NR_126174)
Rhizobium sp. SCAU231 (Leucaena leucocephala, China) (HQ538623)
Sinorhizobium sp. RAM11 (thermal bath, Hungary) (LN651200)
Rhizobium sp. KAR49 (potato, Estonia) (KR055010)
Rhizobium sp. PF-M (Kartchner Caverns, USA) (DQ202284)
49
Rhizobium endolithicum JC140 (sediment sample, India) (NR 133842)
Rhizobium sp. CY2D15 (Ageratina adenophora rhizosphere, China) (HQ231930)
Rhizobium leucaenae LMG 9517T (Phaseolus vulgaris, Brazil) (NR 118993)
Rhizobium leucaenae CCGE 523 (Phaseolus vulgaris, Mexico) (JF318176)
53
Rhizobium leucaenae BR 828 (Leucaena leucocephala, Brazil) (JF318172)
Rhizobium leucaenae BR 10043 (Phaseolus vulgaris, Brazil) (JF318173)
88
Rhizobium tropici P4-6 16S (Phaseolus vulgaris, Portugal) (HM852122)
Rhizobium leucaenae CPAO 29.8 (Phaseolus vulgaris, Brazil) (LMZF01000103)
Rhizobium leucaenae USDA 9039 (Phaseolus vulgaris,Brazil) (AUFB01000074)
Rhizobium tropici 77 (Phaseolus vulgaris, Brazil) (EU488745)
Mesorhizobium sp. IHB B 2272 (rhizoplane of tea roots, India) (HQ686042)
69
Ochrobactrum sp. MI9.3 (bioreactor, Canada) (EU165529)
66
Ochrobactrum sp. 906B6 12AESBL (biogas plant, Germany) (KJ831424)
48
Rhizobium sp. 7A-195 (a mine, Portugal) (KF441633)
Agrobacterium tumefaciens Zutra F/1 (Dahlia sp., Israel) (AJ389909)
94
Rhizobium leguminosarum Mad-7 (Ononis tridentata, Spain) (EF364384)
Rhizobium leguminosarum BKBCF51 (Caragana frutex, Russia) (HQ540584)
Rhizobium sp. LCK33 (Lotus creticus nodules, Tunisia) (KU753919)
55
Rhizobium galegae CCNWSX1209 (Glycine soja nodules, China) (JX524433)
Rhizobium huautlense HNSQ3 (Sophora alopecuroides nodules, China) (JQ821378)
98
Rhizobium sp. HNSQ6 (Sophora alopecuroides nodules, China) (JQ821381)
Rhizobium sp. CCNWNX0072 (Ammopiptanthus Mongolicus, China) (EU697965)
Rhizobium sp. LAR-14 (Arachis hypogaea, Morocco) (EF638791)
0.01


64
Rhizobium leguminosarum bv. viciae USDA 2508 (Vicia faba, Mongolia) (U89831)
Rhizobium leguminosarum Mad-6 (Ononis Tridentata, Spain) (EF364383)
MLSC04 (Lens culimaris, Spain) (MH345075)
MLS05 (Lens culimaris, Spain) (MH345072)
PEPV08 (Phaseolus vulgaris, Spain) (MH345076)
Rhizobium leguminosarum Vaf-108 (Vavilovia formosa, Russia) (CP018228)
Rhizobium sp. L32C558A00 (legume nodules, California) (JX292372)
54
Rhizobium leguminosarum bv. viciae GU-2.2 (Pisum sativum, Toledo) (GQ899184)
Rhizobium leguminosarum CCBAU 85025 (Vicia sativa, Tibet) (EU256423)
100
Rhizobium leguminosarum bv. viciae 248 (Fava bean, East Anglia) (NZ ARRT01000006)
89
Rhizobium leguminosarum bv. viciae USDA 2370T (JQ085246)
Rhizobium laguerreae FB206T (JN558651)
51
Rhizobium leguminosarum bv. viciae UPM1131 (Pisum nodules, Italy) (ATYZ01000039)
100
Rhizobium leguminosarum Mad-7 (Ononis tridentata, Spain) (EF364384)
84
Rhizobium leguminosarum Mad-5 (Ononis Tridentata, Spain) (EF364382)
62
Rhizobium leguminosarum BKBCF51 (Caragana frutex, Russia) (HQ540584)
Rhizobium leucaenae BR 828 (Leucaena leucocephala, Brazil) (JF318172)
Rhizobium leucaenae USDA 9039 (Phaseolus vulgaris,Brazil) (AUFB01000074)
Rhizobium leucaenae CPAO 29.8 (Phaseolus vulgaris, Brazil) (LMZF01000103)
100
Rhizobium leucaenae CCGE 523 (Phaseolus vulgaris, Mexico) (JF318176)
Rhizobium leucaenae BR 10043 (Phaseolus vulgaris, Brazil) (JF318173)
Rhizobium tropici P4-6 16S (Phaseolus vulgaris, Portugal) (HM852122)
Rhizobium leucaenae LMG 9517T (Phaseolus vulgaris, Brazil) (NR 118993)
Rhizobium tropici 77 (Phaseolus vulgaris, Brazil) (EU488745
97
Rhizobium sp. 47.4 (Desmanthus illinoensis, USA) (DQ499526)
99
Rhizobium sp. CY2D15 (Ageratina adenophora rhizosphere, China) (HQ231930)
100
Parahizobium giardinii H152T (NR 026059)
50
Rhizobium sp. Alm-6 (Ononis Tridentata, Spain) (EF364377)
Mesorhizobium sp. IHB B 2272 (rhizoplane of tea roots, India) (HQ686042)
100
Mesorhizobium loti NZP 2213T (NR 025837)
77
Sinorhizobium sp. RAM11 (thermal bath, Hungary) (LN651200)
Ensifer adhaerens LMG 20216T (AM181733)
100
Ochrobactrum sp. 906B6 12AESBL (biogas plant, Germany) (KJ831424)
Ochrobactrum intermedium NBRC 15820T (NR 113812)
97
Ochrobactrum sp. MI9.3 (bioreactor, Canada) (EU165529)
100
Ochrobactrum grignonense NBRC 102586T (NR 114149)
99
Allorhizobium undicola ORS 992T (NR_026463)
98
Rhizobium sp. 7A-195 (a mine, Portugal) (KF441633)
Agrobacterium tumefaciens Zutra F/1 (Dahlia sp., Israel) (AJ389909)
100
Agrobacterium tumefaciens NCPPB 2437T (D14500)
62
Rhizobium endolithicum JC140 (sediment sample, India) (NR_133842)
Rhizobium lemnae L6-16 (Lemna aequinoctialis, Thailand) (NR 126174)
64
99
Rhizobium sp. KAR49 (potato, Estonia) (KR055010) (KR055010)
51
Rhizobium sp. PF-M (Kartchner Caverns, USA) (DQ202284)
Rhizobium sp. SCAU231 (Leucaena leucocephala, China) (HQ538623)
98
67
Neohizobium huautlense SO2T (NR 024863)
Neorhizobium galegae HAMBI540T (HG938353)
Rhizobium sp. LCK33 (Lotus creticus nodules, Tunisia) (KU753919)
70
Rhizobium sp. LAR-14 (Arachis hypogaea, Morocco) (EF638791)
Rhizobium galegae CCNWSX1209 (Glycine soja nodules, China) (JX524433) (EU697965)
74
Rhizobium sp. HNSQ6 (Sophora alopecuroides nodules, China) (JQ821378)
66
Rhizobium huautlense HNSQ3 (Sophora alopecuroides nodules, China) (JQ821381)
Rhizobium sp. CCNWNX0072 (Ammopiptanthus Mongolicus, China) (JX524433)
0.005


MLSC02 (MH345064, MH345046)
PEPV40 (KY624418, KY624419)
rrs IV
88
PEPV31 (MH345070, MH345052)
rrs III
Rhizobium laguerreae FB206T (JN558681, JN558661) (Vicia faba, Tunisia)
97
Rhizobium leguminosarum TLR10 (KC679495, KC679550) (Lens culinaris, Turkey)
A
Rhizobium leguminosarum TLR2 (KC679488, KC679545) (Lens culinaris, Turkey)
56
Rhizobium leguminosarum SLR6 (KC679504, KC679559) (Lens culinaris, Syria)
59
PEPV11 (MH345069, MH345051)
rrs I
. Rhizobium leguminosarum TLR7 (KC679493, KC679548) (Lens culinaris, Turkey)
Rhizobium leguminosarum SLR8 (KC679506, KC679561) (Lens culinaris, Syria)
B
54
74
PEPV37 (MH345071, MH345053)
rrs III
Rhizobium leguminosarum TLR9 (KC679494, KC679549) (Lens culinaris, Turkey)
100
MLS17 (MH345067, MH345049)
rrs III
Rhizobium leguminosarum TLR14 (KC679498, KC679553) (Lens culinaris, Turkey)
PEPV16 (KU196777, KU196776)
rrs II
C
MLS05 (MH345066, MH345048)
MLSC04 (MH345065, MH345047)
rrs V
99
100
PEPV08 (MH345068, MH345050)
Rhizobium leguminosarum SLR7 (KC679505, KC679560) (Lens culinaris, Syria)
D
60
97
Rhizobium leguminosarum SLR4 (KC679502, KC679557) (Lens culinaris, Syria)
Rhizobium sophorae CCBAU 03386T (KJ831252, KJ831235) (Sophora flavescens, China)
Rhizobium leguminosarum USDA 2370T (AJ294376, AJ294405) (Pisum sativum, USA)
100
Rhizobium indigoferae CCBAU 71042T (EF027965, GU552925) (Indigofera sp., China)
Rhizobium ecuadorense CNPSO 671T (NZ LFIO01000381, NZ LFIO01001065) (Phaseolus vulgaris, Ecuador)
Rhizobium acidisoli FH13T (KJ921098, KJ921069) (Phaseolus vulgaris, Mexico)
61
Rhizobium hidalgonense FH14T (KJ921099, KJ921070) (Phaseolus vulgaris, Mexico)
74
100
. Rhizobium leguminosarum NH05 (KJ921109, KJ921077) (Phaseolus vulgaris, Mexico)
Rhizobium anhuiense CCBAU 23252T (KF111980, KF111890) (Vicia faba, China)
Rhizobium fabae CCBAU 33202T (EF579941, EF579929) (Vicia faba, China)
100
Rhizobium pisi DSM 30132T (EF113134, EF113149) (Pisum sativum, unknown origin)
Rhizobium sophoriradicis CCBAU 03470T (KJ831248, KJ831231) (Sophora flavescens, China)
100
Rhizobium etli IE4803 (CP007641) (Phaseolus vulgaris, Mexico)
Rhizobium aegyptiacum 1010T (KU664569, KU664561) (Trifolium alexandrinum, Egypt)
Rhizobium lentis BLR27T (JN649031, JN648941) (Lens culinaris, India)
80
100
Rhizobium etli CFN42T (NC_007761) (Phaseolus vulgaris, Mexico)
74
Rhizobium etli NXC12 (CP020906) (Phaseolus vulgaris, Mexico)
92
56
Rhizobium etli N561 (CP013500) (Phaseolus vulgaris, Mexico)
Rhizobium bangladeshense BLR175T (JN649057, JN648967) (Lens culinaris, India)
Rhizobium sp. CIAT894 (CP020947) (Phaseolus vulgaris, Colombia)
Rhizobium esperanzae CNPSo 668T (NZ MXPU01000020, NZ MXPU01000002) (Phaseolus vulgaris, Mexico)
Rhizobium binae BLR195T (JN649058, JN648968) (Lens culinaris, India)
Rhizobium leguminosarum CCGM1 (JFGP01000004) (Phaseolus vulgaris, Mexico)
100
Rhizobium phaseoli ATCC 14482T (EF113136, EF113151) (Phaseolus vulgaris, USA)
98
Rhizobium phaseoli Brasil 5 (CP020896) (Phaseolus vulgaris, Brazil)
0.01


Rhizobium leguminosarum TLR2 (KC679641) (*Lens culinaris*, Turkey)
Rhizobium leguminosarum TLR7 (KC679644) (*Lens culinaris*, Turkey)
Rhizobium fabae CCBAU 33202^{T} (N580683) (*Vicia faba*, China)
MLS17 (MH345057)
Rhizobium binae BLR195^{T} (JN649012) (*Lens culinaris*, India)
Rhizobium laguerreae FB206^{T} (KC608575) (*Vicia faba*, Tunisia)
Rhizobium leguminosarum SLR8 (KC679656) (*Lens culinaris*, Syria)
PEPV37 (MH345060)
PEPV40 (MH345062)
MLS05 (MH345054)
Rhizobium leguminosarum TLR10 (KC679647) (*Lens culinaris*, Turkey)
Rhizobium leguminosarum TLR14 (KC679650) (*Lens culinaris*, Turkey)
PEPV31 (MH345063)
Rhizobium leguminosarum TLR9 (KC679646) (*Lens culinaris*, Turkey)
Rhizobium leguminosarum USDA 2370^{T} (FJ596038) (*Pisum sativum*, USA)
Rhizobium pisi DSM 30132^{T} (HQ394248) (*Pisum sativum*, unknown origin)
Rhizobium leguminosarum SLR4 (KC679653) (*Lens culinaris*, Syria)
Rhizobium lentis BLR105 (JN649004) (*Lens culinaris*, India)
Rhizobium bangladeshense BLR175^{T} (JN649011) (*Lens culinaris*, India)
MLSC02 (MH345055)
PEPV16 (MH345061)
MLSC04 (MH345056)
PEPV08 (MH345058)
PEPV11 (MH345059)
Rhizobium hidalgonense FH14^{T} (NZ NWSY01000044) (*Phaseolus vulgaris*, Mexico)
Rhizobium leguminosarum CCGM1 (JFGP01000014) (*Phaseolus vulgaris*, Mexico)
Rhizobium phaseoli ATCC 14482^{T} (HM441255) (*Phaseolus vulgaris*, USA)
Rhizobium etli CFN42^{T} (NC_007761) (*Phaseolus vulgaris*, Mexico)
Rhizobium etli IE4803 (CP007643) (*Phaseolus vulgaris*, Mexico)
Rhizobium esperanzae CNPSo 668^{T} (NZ MXPU01000033) (*Phaseolus vulgaris*, Mexico)
Rhizobium acidisoli FH23^{T} (NZ NWSS01000043) (*Phaseolus vulgaris*, Mexico)
Rhizobium etli N561 (CP013502) (*Phaseolus vulgaris*, Mexico)
Rhizobium ecuadorense CNPSO 671^{T} (NZ_LFIO01000036) (*Phaseolus vulgaris*, Ecuador)
Rhizobium phaseoli Brasil 5 (CP020898) (*Phaseolus vulgaris*, Brazil)
Rhizobium sp. CIAT894 (CP020950) (*Phaseolus vulgaris*, Colombia)
94
97
72
65
72
58
91
99
99
61
62
99
62
100
88
100
99
100
95
51
93
91
1
2
3
sv viciae
sv phaseoli
0.05

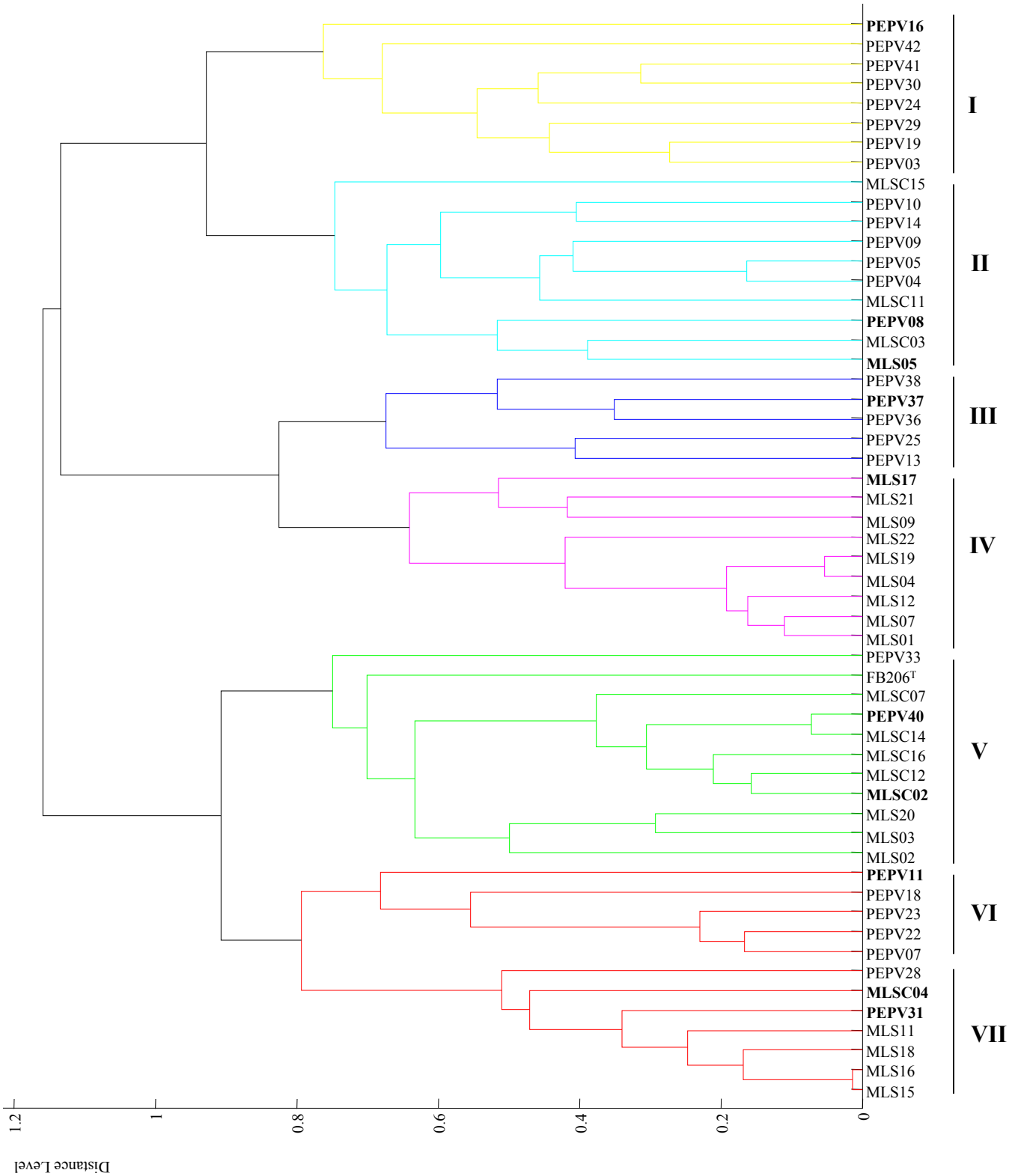


Figure S1. Cluster analysis of MALDI-TOF MS spectra of strains isolated in this study. Distance is displayed in relative units. Representative strains of each group are marked in bold.

Table 1. Results of MALDI-TOF MS analysis of strains analysed in this study.

Strains	MALDI-TOF MS group	Organism (best match)	Score value
Isolated from *P. vulgaris*			
PEPV03	I	*Rhizobium laguerreae* FB206^T	2.131
PEPV04	II	*Rhizobium laguerreae* FB206^T	2.200
PEPV05	II	*Rhizobium laguerreae* FB206^T	2.215
PEPV07	VI	*Rhizobium laguerreae* FB206^T	2.131
PEPV08	II	*Rhizobium laguerreae* FB206^T	2.201
PEPV09	II	*Rhizobium laguerreae* FB206^T	2.244
PEPV10	II	*Rhizobium laguerreae* FB206^T	2.227
PEPV11	VI	*Rhizobium laguerreae* FB206^T	2.080
PEPV13	III	*Rhizobium laguerreae* FB206^T	2.272
PEPV14	II	*Rhizobium laguerreae* FB206^T	2.314
PEPV16	I	*Rhizobium laguerreae* FB206^T	2.181
PEPV18	VI	*Rhizobium laguerreae* FB206^T	2.069
PEPV19	I	*Rhizobium laguerreae* FB206^T	2.181
PEPV22	VI	*Rhizobium laguerreae* FB206^T	2.069
PEPV23	VI	*Rhizobium laguerreae* FB206^T	2.123
PEPV24	I	*Rhizobium laguerreae* FB206^T	2.083
PEPV25	III	*Rhizobium laguerreae* FB206^T	2.259
PEPV28	VII	*Rhizobium laguerreae* FB206^T	2.187
PEPV29	I	*Rhizobium laguerreae* FB206^T	2.050
PEPV30	I	*Rhizobium laguerreae* FB206^T	2.221
PEPV31	VII	*Rhizobium laguerreae* FB206^T	2.131
PEPV33	V	*Rhizobium laguerreae* FB206^T	2.297
PEPV36	III	*Rhizobium laguerreae* FB206^T	2.131
PEPV37	III	*Rhizobium laguerreae* FB206^T	2.136
PEPV38	III	*Rhizobium laguerreae* FB206^T	2.228
PEPV40	V	*Rhizobium laguerreae* FB206^T	2.317
PEPV41	I	*Rhizobium laguerreae* FB206^T	2.277
PEPV42	I	*Rhizobium laguerreae* FB206^T	2.157
Isolated from *L. culinaris*			
MLS01	IV	*Rhizobium laguerreae* FB206^T	2.268
MLS02	V	*Rhizobium laguerreae* FB206^T	2.317
MLS03	V	*Rhizobium laguerreae* FB206^T	2.382
MLS04	IV	*Rhizobium laguerreae* FB206^T	2.415
MLS05	II	*Rhizobium laguerreae* FB206^T	2.091
MLS07	IV	*Rhizobium laguerreae* FB206^T	2.351
MLS09	IV	*Rhizobium laguerreae* FB206^T	2.328
MLS11	VII	*Rhizobium laguerreae* FB206^T	2.151
MLS12	IV	*Rhizobium laguerreae* FB206^T	2.379
MLS15	VII	*Rhizobium laguerreae* FB206^T	2.173
MLS16	VII	*Rhizobium laguerreae* FB206^T	2.166
MLS17	IV	*Rhizobium laguerreae* FB206^T	2.454
MLS18	VII	*Rhizobium laguerreae* FB206^T	2.198
MLS19	IV	*Rhizobium laguerreae* FB206^T	2.397
MLS20	V	*Rhizobium laguerreae* FB206^T	2.259
MLS21	IV	*Rhizobium laguerreae* FB206^T	2.418
MLS22	IV	*Rhizobium laguerreae* FB206^T	2.300
MLSC02	V	*Rhizobium laguerreae* FB206^T	2.270
MLSC03	II	*Rhizobium laguerreae* FB206^T	2.071
MLSC04	VII	*Rhizobium laguerreae* FB206^T	2.186
MLSC07	V	*Rhizobium laguerreae* FB206^T	2.281
MLSC11	II	*Rhizobium laguerreae* FB206^T	2.151
MLSC12	V	*Rhizobium laguerreae* FB206^T	2.317
MLSC14	V	*Rhizobium laguerreae* FB206^T	2.189
MLSC15	II	*Rhizobium laguerreae* FB206^T	2.221
MLSC16	V	*Rhizobium laguerreae* FB206^T	2.266